



**QUEEN'S
UNIVERSITY
BELFAST**

Imitate or innovate? Children's innovation is influenced by the efficacy of observed behaviour

Carr, K., Kendal, R. L., & Flynn, E. G. (2015). Imitate or innovate? Children's innovation is influenced by the efficacy of observed behaviour. *Cognition*, 142, 322-332. <https://doi.org/10.1016/j.cognition.2015.05.005>

Published in:
Cognition

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2015 The Authors. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)
This is an open access article published under a Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.



Imitate or innovate? Children's innovation is influenced by the efficacy of observed behaviour



Kayleigh Carr^{a,*}, Rachel L. Kendal^b, Emma G. Flynn^c

^a Department of Psychology, Centre for the Coevolution of Biology and Culture, Durham University, UK

^b Department of Anthropology, Centre for the Coevolution of Biology and Culture, Durham University, UK

^c School of Education, Centre for the Coevolution of Biology and Culture, Durham University, UK

ARTICLE INFO

Article history:

Received 23 December 2014

Revised 5 May 2015

Accepted 6 May 2015

Available online 11 June 2015

Keywords:

Innovation

Behaviour efficacy

Imitation

Selective social learning

Asocial learning

Trade-offs

ABSTRACT

This study investigated the age at which children judge it futile to imitate unreliable information, in the form of a visibly ineffective demonstrated solution, and deviate to produce novel solutions ('innovations'). Children aged 4–9 years were presented with a novel puzzle box, the Multiple-Methods Box (MMB), which offered multiple innovation opportunities to extract a reward using different tools, access points and exits. 209 children were assigned to conditions in which eight social demonstrations of a reward retrieval method were provided; each condition differed incrementally in terms of the method's efficacy (0%, 25%, 75%, and 100% success at extracting the reward). An additional 47 children were assigned to a no-demonstration control condition. Innovative reward extractions from the MMB increased with decreasing efficacy of the demonstrated method. However, imitation remained a widely used strategy irrespective of the efficacy of the method being reproduced (90% of children produced at least one imitative attempt, and imitated on an average of 4.9 out of 8 attempt trials). Children were more likely to innovate in relation to the tool than exit, even though the latter would have been more effective. Overall, innovation was rare: only 12.4% of children innovated by discovering at least one novel reward exit. Children's prioritisation of social information is consistent with theories of cultural evolution indicating imitation is a prepotent response following observation of behaviour, and that innovation is a rarity; so much so, that even maladaptive behaviour is copied.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Social learning provides the foundation for culture. Acquiring information through observation is a rapid, cheap and largely efficient way to learn. Yet, on occasion, social information is outdated or inappropriate, especially in changing environments; thus its use must be modulated to support accurate and reliable information acquisition (Boyd & Richerson, 1985; Kameda & Nakanishi, 2002). Accordingly, personal sampling of the environment, even if costly, is a necessity (Laland, 2004). Theoretical models have suggested many learning heuristics (cultural transmission biases; Boyd & Richerson, 1985 and social learning strategies; Laland, 2004) which enable selectivity in social learning. These heuristics help direct whom, when and what we copy by inducing accuracy-cost evaluations of observed and personal information and, in turn, adaptive

trade-offs in reliance on social and asocial (individual) learning (Kendal, Coolen, & Laland, 2009; Kendal, Coolen, van Bergen, & Laland, 2005).

Adaptive informational trade-offs have been shown in a variety of non-human animals (including species of fish, rats, monkeys and birds; see Galef & Laland, 2005; Kendal et al., 2009). By pitting social and personal information against one another, it appears that, 'animals use social information primarily as plan B, or a backup when personal information is too costly to obtain, unreliable or outdated' (Rieucou & Giraldeau, 2011, p. 950). In van Bergen, Coolen, and Laland (2004), three groups of nine-spined stickleback fish were provided with personal information that varied in its level of reliability (56%, 78% or 100% reliable). This information related to the profitability of food patches within the experimental tank, and was determined by the number of trials in which 'rich' and 'poor' feeders could be accessed. A social ('public') demonstration then provided conflicting information as to the location of the rich feeder. In spite of this demonstration, a significant number of sticklebacks within the 100% group (19 of 23) continued to visit the feeder they had personally experienced as rich,

* Corresponding author at: Department of Psychology, Durham University, South Road, Durham DH1 3LE, UK.

E-mail addresses: kayleigh.carr@durham.ac.uk (K. Carr), rachel.kendal@durham.ac.uk (R.L. Kendal), e.g.flynn@durham.ac.uk (E.G. Flynn).

thus negating the conflicting social information. As with van Bergen, Coolen, and Laland (2004), in the current study we manipulated information reliability with the aim of observing adaptive trade-offs in learning. However, given children's proclivity for imitation, and apparent tendency to collect social information despite possessing adequate personal information (Wood, Kendal, & Flynn, 2013a), we did so by manipulating the reliability of social information.

Children are exceptional imitators from a young age, reproducing behaviour with high levels of fidelity across contexts (Matheson, Moore, & Akhtar, 2013) and in the absence of causal knowledge of its relevancy (Lyons, Young, & Keil, 2007). Indeed, they are deemed 'cultural magnets' (Flynn, 2008) in their ability to both rapidly acquire and transmit information socially (Flynn & Whiten, 2008; Hopper, Flynn, Wood, & Whiten, 2010). However, children are not blind to the quality of information they observe. By altering the frequency and fidelity with which they imitate, in line with the perceived goal of a demonstration (Bekkering, Wohlschläger, & Gattis, 2000; Carpenter, Call, & Tomasello, 2005), model reliability and intentionality (Birch, Vauthier, & Bloom, 2008; Carpenter, Akhtar, & Tomasello, 1998), task difficulty and prior experience (Gardiner, Bjorklund, Greif, & Gray, 2012; Pinkham & Jaswal, 2011; Williamson & Meltzoff, 2011; Wood et al., 2013a), children display rationality and flexibility in their social learning (Koenig & Sabbagh, 2013; Mills, 2013; Over & Carpenter, 2012).

A variety of factors, including context, model characteristics and information content, affect the use of social information (Rendell et al., 2011; Wood, Kendal, & Flynn, 2013b); here, our focus is on the efficacy of the information content. Action efficacy should arguably be a foremost determinant of what (and if) we choose to copy. By 3 years of age children distinguish correct from incorrect actions in their imitative behaviour, only reproducing those that have a desired causal effect (Want & Harris, 2001). Further, prior personal difficulty with a task does not induce 3-year-olds to have a copy-all approach when non-efficacious acts are demonstrated (Williamson, Meltzoff, & Markman, 2008). If a causal relationship is unknown, faithful imitation may result. However, if action sequences are repeatedly poor at producing desired outcomes, their efficacy should be questioned and imitation less likely to occur. Thus, logically, in circumstances under which a sequence of behaviour is never or rarely effective at achieving a goal, individuals should try new methods.

Few studies have attempted to examine how evaluations of efficacy affect selective imitation, and subsequent novel action production (or *innovation*). Schulz, Hooppell, and Jenkins (2008) tested 18-month-olds and 4-year-olds in conditions that differed in an action's efficacy: deterministic, in which the actions activated the toy on all trials and stochastic, in which actions activated the toy on 50% of trials. Children of both age groups imitated with significantly lower fidelity in the stochastic condition than the deterministic condition, irrespective of whether the action satisfied the explicitly stated goal of the model. Thus, in the stochastic condition, efficacy overrides pedagogy. However, as Schulz et al. (2008) acknowledge, the potential for alternative responses on the task, and the opportunity to observe behavioural innovation, was limited.

In recent years, interest in childhood innovation has grown, and studies suggest that, in the tool-use domain, innovation is a relatively late-developing capacity (Beck, Apperly, Chappell, Guthrie, & Cutting, 2011; Hanus, Mendes, Tennie, & Call, 2011; Nielsen, 2013) and a rare response for children (Whiten & Flynn, 2010). Factors such as functional fixedness (German & Defeyter, 2000), explicit instruction (Bonawitz et al., 2011), prior social information (Wood et al., 2013a), and task structure (Cutting, Apperly, Chappell, & Beck, 2014) likely constrain it. Innovation can be

delineated in terms of arising from asocial learning (innovation by independent invention) or a combination of asocial and social learning (innovation by modification; Carr, Kendal, & Flynn, under revision). Most research investigating children's innovation has examined novel tool invention as opposed to novel modification. Yet, examination of the latter is critical as it is of great importance for cumulative culture (Lewis & Laland, 2012), where, over generations, humans build upon and improve pre-existing knowledge (Dean, Kendal, Schapiro, Thierry, & Laland, 2012). Currently we do not know whether innovation by modification has the same late developmental trajectory as independent invention. The current study addresses this issue through the provision of social demonstrations to individual children, across the age range of 4–9 years, followed by multiple response trials, thus providing many opportunities for innovation as well as multiple tools with which to innovate.

We ask, when evaluating efficacy of observed actions, at which point do children judge it futile to imitate? Do we see different assessments of redundancy at different ages? And does varying action efficacy make children more likely to innovate (produce novel behaviour) when given sufficient opportunity and means to do so? Even if children do not know of a behavioural alternative, they should nevertheless explore novel actions (Schulz et al., 2008) – trading-off social information for potentially more reliable personal information.

Our study used a novel artificial fruit (Whiten, Custance, Gomez, Teixidor, & Bard, 1996), the Multiple-Methods Box (MMB), a puzzle box offering scope for exploration and innovation (we distinguish exploration and innovation here as they are regarded as qualitatively distinct (Reader & Laland, 2003): you may explore, but you may not always innovate). Drawing from van Bergen et al. (2004), children were provided with social demonstrations that differed in solution efficacy: the proportion of trials (0%, 25%, 75%, 100%) that a reward could be extracted from the exit door of the MMB. Multiple demonstration and attempt trials were provided to reduce the likelihood that the novel task and experimental context would incite a copy-when-uncertain bias (Laland, 2004) and to monitor if, and how, participants changed their reliance on social and/or personal information over trials (Flynn & Smith, 2012; Wood et al., 2013a). With increasing experience with the MMB, both through observation and personal use, participants could establish the demonstrated method's efficacy and, in the lower efficacy conditions, appreciate the redundancy of repeating a method that simply did not work.

Children aged 4–9 years were selected so as to capture developmental change and is in keeping with that of previous innovation research (Beck et al., 2011). Moreover, children are adept at assessing efficacy by 4 years (Want & Harris, 2001; Williamson et al., 2008) and able to differentiate information that is reliable 75% of the time from information that is reliable 25% of the time (Pasquini, Corriveau, Koenig, & Harris, 2007). We predicted, in line with Want and Harris (2001) and Schulz et al. (2008), that lower levels of solution efficacy would be associated with reduced imitation (lowered fidelity to the socially demonstrated method), and, further, increased innovation (specifically, innovations that altered the reward exit and thus allowed for extraction). Moreover, we anticipated that older children would be better equipped to both evaluate levels of solution efficacy (resulting in a stronger negative relationship between efficacy and innovation with increasing age) and reach effective innovative solutions (with the greatest rates of successful innovation being seen in the oldest age group). In turn, we predicted that, overall, the oldest children would be the least faithful to the socially demonstrated method. Finally, given the range of novel behaviours that could be produced with the MMB, we explore how participants deviated from the socially demonstrated method (if and when they did) with regard to whether they

changed the tool, access point or, most effectively, the exit. We assessed the children's performance against the performance of adults, whom we predicted should innovate, particularly in the lowest efficacy condition.

2. Method

2.1. Participants

Two hundred and fifty-six children (128 males) from three primary schools in the North East of England participated. Three age groups were created: 4–5 years ($N = 73$, $M = 5$ years 4 months (5;4), range 4;8–5;11), 6–7 years ($N = 96$, $M = 7;0$, range 6;0–7;10), and 8–9 years ($N = 87$, $M = 8;10$, range 8;0–9;9). Forty-five Durham University students also participated (23 male, $M = 20$ years 7 months (20;7), range 18;6–27;7).

2.2. Materials

A novel puzzle box task, the 'Multiple-Methods Box' (MMB; see Fig. 1), was used. The MMB contains two levels separated by an opaque platform. The upper transparent level featured: an entry chute for a reward (a capsule containing a sticker which was inserted by the experimenter); four entrances, one of which required the rotation of a dial for access and three of which could also function as reward extraction points; and a small circular hole in the platform floor. If the capsule was manipulated to fall through this hole (as in the social demonstrations) it dropped to a lower opaque level of the MMB via a concealed slope to rest behind a blue exit door. A small independent remote control device was used to discretely lock and unlock the exit door in line with predetermined levels of solution efficacy. When unlocked the door could be lifted to acquire the capsule from behind.

Three tools were available: a fork, a hook and a sweep tool (Fig. 1b). The varying dimensions of both the MMB and the tools introduced an additional problem solving component to the task by limiting random application of the tools; that is, not all tools fitted into all access points or were long enough to manipulate the capsule to all exit holes. Further, the fork and sweep tool could be joined and used in combination to extract the reward across a longer distance than the other single tools. The social demonstration involved inserting the fork tool into the smaller inverted

T-shaped entrance (labelled 1 in Fig. 1), the reward was caught in the 'U' of the fork and manoeuvred so that it fell through the hole in the platform floor.

2.3. Design

Children from each age group were randomly allocated to one of four social experimental conditions, differing incrementally in the efficacy of the demonstrated method of reward extraction. The method itself was consistent across all demonstrations and conditions. Method efficacy was operationalised as the number of demonstration trials, out of eight, in which the capsule could be removed from the exit door. The method was efficacious on either 0 of 8 trials (0% condition, $N = 60$), 2 of 8 trials (25% condition, $N = 48$), 6 of 8 trials (75% condition, $N = 50$) or 8 of 8 trials (100% condition, $N = 51$). Importantly, the level of method efficacy observed during the experimenter's demonstrations was mirrored in participants' own subsequent attempts with the task, such that their personal experience with the MMB matched their observational experience (if they chose to reproduce the demonstrated behaviour). A further 47 children were assigned to a no-demonstration control condition in which they witnessed no social demonstrations (see Table 1 for the distribution of participants across groups). This condition provided a baseline measure of performance on the task, specifically the level of performance of the actions presented within the social demonstration and the level of new method generation without prior method demonstration. The adult participants were allocated to either the 0% or 75% efficacy condition as it was here that major differences were seen in the child sample.

2.4. Procedure

Children were tested individually in a quiet area of their school. First, they were familiarised with the MMB during a short warm-up phase. To attempt to reduce assumed experimenter expertise and potential model-based biases (Wood, Kendal, & Flynn, 2012), the box was proclaimed as belonging to a friend, "This is actually my friend's box, and my friend told me that when this egg [the capsule] goes into the box you have to try and get it out. Inside this egg is a sticker. If you get it out of the box, we can start a sticker pile for you and we'll see how many you can

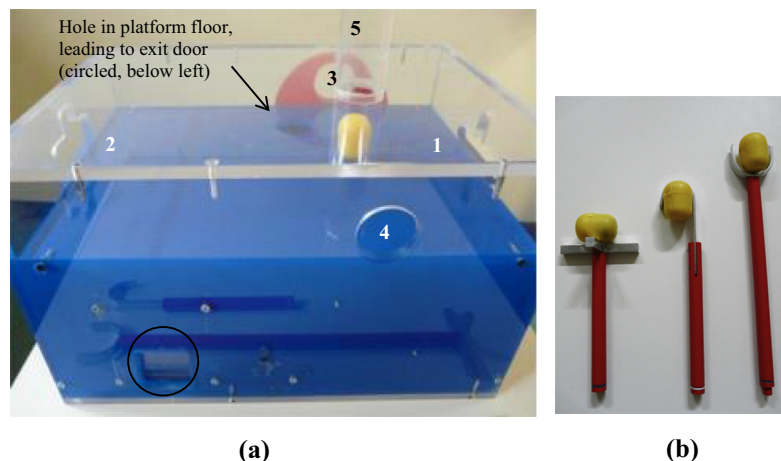


Fig. 1. The Multiple-Methods Box (MMB) and associated tools. (a) Access points labelled 1–5: (1) 'Social', small inverted T-shape, used in social demonstrations, (2) 'End', large inverted T-shape, opposite 'Social', (3) 'Dial', circular hole, revealed by aligning the circle of a dial with a circle in the side of the box, (4) 'Dial Opposite', and (5) 'Entry Chute', a circular hole into which the reward was dropped. (b) Three tools were available, from right to left: fork, hook and sweep. The position of the capsule in relation to each tool demonstrates the main method of manoeuvre. The fork and sweep tool could be joined and used in combination to extract the reward, with the extra length affording extraction across the full length of the MMB, and can be seen in the reflection at the base of the box (a).

Table 1

Distribution of participants across the experimental conditions and age groups.

	Baseline	0%	25%	75%	100%	Total
4–5 yrs	14(8)	17(9)	14(7)	15(9)	13(8)	73(41)
6–7 yrs	18(9)	20(11)	18(10)	19(9)	21(11)	96(50)
8–9 yrs	15(8)	23(10)	16(7)	16(6)	17(6)	87(37)
Total	47(25)	60(30)	48(24)	50(24)	51(25)	256(128)

Note. Number of males given in parentheses.

get”. Anecdotally, many children appeared to accept this premise by enquiring into the name of the friend. The tools were presented alongside the box: “Can you see these tools here? My friend also told me that some of these tools can be joined together”.

Children in the no-demonstration control condition received a prompt to begin interacting with the MMB immediately following this familiarisation: “You can have some turns at seeing if you can get the egg out of the box. You can do anything you like.” The exit door was unlocked throughout for control participants, and up to five tool insertions permitted per attempt. Children in the social conditions were informed: “I’m going to have eight turns at trying to get the egg out of the box. Let’s see if it works”. The experimenter proceeded to demonstrate the same method of reward retrieval (fork tool through ‘Social’ access point, capsule to exit door via hole in floor) eight times with only the outcome differing between the four experimental groups. Neutral comments, “I got it out of the box/I didn’t get it out of the box”, were made after each demonstration. As the concealed exit chute connecting the circular hole in the upper platform floor and the lower exit door was capable of holding eight capsules, it was not necessary to remove ‘locked’ capsules in between experimenter demonstrations. However, for those conditions in which locked capsules had to be removed prior to participants’ attempts (0–75% conditions), children were distracted with a non-cognitively demanding task (organising sheets of stickers) for the very short time (<10 s) it took to remove these capsules.

Participants were given a maximum of eight attempt trials, over a period of five minutes (if the eight trials were not completed within this time, which was rare, testing ceased). Participants who had received social demonstrations were told, “Now it’s your turn to see if you can get the egg out of the box. You can do anything you like”. Each trial constituted one participant’s attempt, for which strict criteria were applied. An attempt was defined as the insertion of a tool into the box with the purposeful intention, or realisation, of making contact with the capsule prior to the tool’s extraction. ‘Purposeful’ denotes engagement with the task as indicated by head and gaze orientation and ‘intention’ evidenced when a tool was fully inserted but too short to reach the capsule. An attempt was complete when a tool was fully extracted (even if then replaced into the same access point). Some innovative methods of reward retrieval involved performing more than one action – for example, pushing the capsule with the fork tool towards the ‘End’ of the MMB before using the hook tool to extract it. In the event that a child displayed continued purposeful intentionality and interaction with the MMB, therefore, this was considered part of the same attempt. The apparatus was re-baited upon commencing each trial, unless full contact with the capsule was not previously made or the capsule was moved only a very small distance. The removal of the lid of the box, concealed by a large fabric sheet, allowed capsules to be quickly retrieved in the event of their unsuccessful extraction. As with demonstrations, neutral comments were made following each attempt trial (“You got it out of the box/You didn’t get it out of the box”).

For comparability, and to control for primacy and recency effects, the demonstration sequence of the two conditions

involving uncertainty (25% and 75%) began and ended with a success (S, door unlocked) followed by an unsuccess (U, door locked). The full demonstration sequence for the 25% condition was thus: S, U, U, U, U, U, S, U, and for the 75% condition: S, U, S, S, S, S, S, U. The same sequences were implemented for participants’ subsequent attempts with the MMB. In this attempts phase, the experimenter ensured only one capsule was extracted on those occasions in which the exit door was unlocked and additional capsules had accumulated in the exit chute. Whilst recognising that it would not always be feasible to fully mirror the efficacy of demonstrated social information in participants’ attempts, given that different numbers of the socially demonstrated method could be attempted prior to the enactment of alternative methods, at the very least participants were given some experience of efficacy variability in their first two trials (i.e., success followed by unsuccess) for these two conditions. It should be noted that enactment of alternative methods that utilised the exit door (alternative by way of a novel tool and/or access point) resulted in the same experience of efficacy as that of the socially demonstrated method.

At the end of testing all children were praised for their performance and rewarded with a sticker irrespective of their level of success (small stickers collected during testing were traded for one larger and more desirable sticker). The above protocol was followed for adult participants in a University laboratory, within either the 0% or 75% conditions. They received departmental credits for their participation or an Amazon voucher, irrespective of their performance.

2.5. Coding and inter-rater reliability

The performance of each participant was scored for a number of variables in each response trial: (a) tool selected, (b) access point used, (c) exit location (if any), (d) outcome (no outcome, capsule to exit door but no extraction, and extraction), and (e) learning strategy. Full rationales for the different strategies are presented in Section 3 but, in short, the strategy was determined by the aforementioned (a)–(c), such that:

- Imitation = same tool, same access point, and same exit as used in social demonstrations.
- Tool/access point innovation = different tool and/or access point, but same exit as used in social demonstrations.¹
- Exit innovation = different or same tool/access point, and different exit as used in social demonstrations (unlike alterations to the tool or access point, discovering a new exit has the potential to change the outcome of the task).
- Unsuccessful action = abandoned attempt prior to removal of capsule or it reaching the exit door.

From these individual response trial variables, several addi-

¹ This is in essence ‘end-state emulation’ (Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009). However, as the end state was manipulated to produce method efficacy, it was not possible to investigate the development of end-state emulation in and of itself.

Table 2
Attempt trial variables subject to statistical analysis.

Variable	Description
Copying fidelity	Score of 1–4 was given for each trial: 1 for no new components (tool/access/exit), 2 for one new component, 3 for two new components, and 4 for three new components. These were summed across the eight attempt trials (max. 32)
Tool/access/exit innovations	Total number of novel (to the child) tools/access points/exits used across attempt trials
Imitation attempts	Total number of attempts in line with strict imitation definition (same tool, access and exit; max. 8)
Tool/access innovation attempts	Total number of attempts in line with tool/access point innovation definition (new tool and/or access, but same exit; max. 8)
Exit innovation attempts	Total number of attempts in line with exit innovation definition (same or different tool and/or access, and different exit; max. 8)
Alternative methods	Total number of different methods (new combinations of tool, access and exit) enacted, excluding socially demonstrated method and irrespective of success (max. 8)
Extractions	Total number of successful capsule extractions, irrespective of extraction method (max. 8)

Note. Attempts at retrieving the capsule were deemed more revealing than successful extractions, as, according to the experimental design of the study, on some trials the capsule reached the exit door but it was locked and so could not be extracted. Here, participants' persistence with an unsuccessful method was evident.

tional variables were created to capture overall task behaviour (Table 2). The experimenter, KC, coded 100% of the sample from video tape. An independent observer, blind to the hypotheses of the study, coded 20% of the sample. All Cohen's Kappa scores and correlation values were 0.85 or above, showing an excellent level of inter-rater reliability.

2.6. Statistical methods

As the data were not normally distributed, non-parametric tests were used. Although we were selective with follow-up tests (Mann–Whitney and Wilcoxon rank-sum), to avoid inflating the Type I error rate a Bonferroni correction was applied by dividing the critical significance level of .05 by the total number of tests conducted. Probability values reported with an asterisk indicate the significance level required to reject the null hypothesis following this correction.

3. Results

The results are presented in four sections. First, we explore how control participants' success and method use compared to that of 100% efficacy social demonstration participants. The 100% condition is the most valid comparison as the door remained unlocked for all trials, as it did in the no-demonstration control condition. The second section considers copying fidelity, broadly defined and then in relation to typical definitions of imitation, followed in the third section by a consideration of deviations from demonstrated behaviour. Finally, innovation, along with its various manifestations, and its role in low efficacy social conditions is investigated. As the sex of participants was not found to significantly affect our outcome measures, it was excluded from all reported analyses. All tests are two-tailed unless otherwise stated.

3.1. What were the level of success and methods used by children in the no-demonstration condition?

Of the 47 controls, nine (19.1%, six males, four 4–5 years and five 6–7 years) failed to produce one attempt; instead, they touched and explored the MMB with their hands but never made

contact with the capsule (despite having the tools introduced at the beginning of their turn). In comparison, all 209 children from the social conditions attempted the task (whether successful or unsuccessful in terms of extraction). 36 of the 47 controls succeeded in making at least one capsule extraction. However, control participants achieved significantly fewer extractions ($Mdn = 5$, $SD = 2.56$) than those in the 100% condition across the attempt trials ($Mdn = 7$, $SD = 2.09$; Mann–Whitney $U = 588.50$, $z = -4.42$, $p < .001$).

The main point of concern which the control condition allowed us to address was whether the socially demonstrated method was a naturally salient response to the task. Of the 38 control participants who produced at least one attempt, only two produced the method of social demonstrations on more attempt trials than any other method. Controls ($Mdn = 0$, $SD = 0.81$) also performed the method of social demonstrations on significantly fewer attempt trials than participants in the 100% condition ($Mdn = 6$, $SD = 2.63$; $U = 131.00$, $z = -8.01$, $p < .001$), whilst attempting a significantly greater number of alternative methods (control: $Mdn = 2.5$, $SD = 1.61$; 100%: $Mdn = 1$, $SD = 1.08$; $U = 361.00$, $z = -5.17$, $p < .001$). Control participants produced a median of 2.5 alternative methods, thus they did not simply discover one means of solving the task and adhere to it. Nevertheless, the majority of children repeated successful methods ($N = 30$, $Mdn = 2$, $SD = 2.42$). Within the control group, the 8- to 9-year-olds produced a greater median number of successful alternative methods ($Mdn = 3$, $SD = 1.55$) than 6- to 7-year-olds ($Mdn = 2$, $SD = 1.61$) and 4- to 5-year-olds ($Mdn = 1$, $SD = 1.70$).

Whilst any successful method discovered in the no-demonstration control condition would technically constitute an innovation, because it is a different kind of innovation to that required in the social conditions (invention versus modification) it is not possible to compare them like-for-like. Hence, the focus above is on alternative methods.

3.2. Did children imitate the socially demonstrated method?

Children received a score according to the number of new components each attempt contained (explained in Table 2). A score of 1 indicated faithful reproduction of the socially demonstrated method, whilst 4 indicated complete deviation from this method. The attempts that had no outcome (they were abandoned, by extracting the tool from the box before an outcome was produced) could receive a maximum score of 3 only due to the unknown exit. A total of 122 participants (58%) produced at least one such abandoned attempt and they accounted for 15.6% of all attempts. The following analyses were run with the abandoned attempts (unsuccessful actions) both included and excluded, with the same effects found. We report the former.

Analysis of total scores, summed across the eight attempt trials, revealed no significant differences between efficacy conditions (Kruskal–Wallis $H(3) = 2.82$, $p = .42$). Children's mean scores in all conditions ranged from 11 to 12.5 (for adults, the overall mean score was larger at 17.42; range of 22). Given that the minimum possible score was 8, denoting complete fidelity throughout attempts, and the maximum 32, children showed little (though some) deviation from demonstrated behaviour, irrespective of condition. Age differences were found in copying fidelity ($H(2) = 12.32$, $p = .002$). Specifically, 4- to 5-year-olds showed significantly higher copying fidelity ($Mdn = 9$, $SD = 3.98$) than 6- to 7-year-olds ($Mdn = 11$, $SD = 4.02$, $U = 1732.5$, $z = -2.53$, $p = .01$) and 8- to 9-year-olds ($Mdn = 11.5$, $SD = 4.32$, $U = 1397$, $z = -3.43$, $p = .001$).

Definitions of imitation usually require that both the goal, and the specific actions used to achieve it, are recognised and reproduced (Tomasello, 1990). Such 'pure' imitation, involving use of the fork tool, through the 'social' access point, and extraction (or

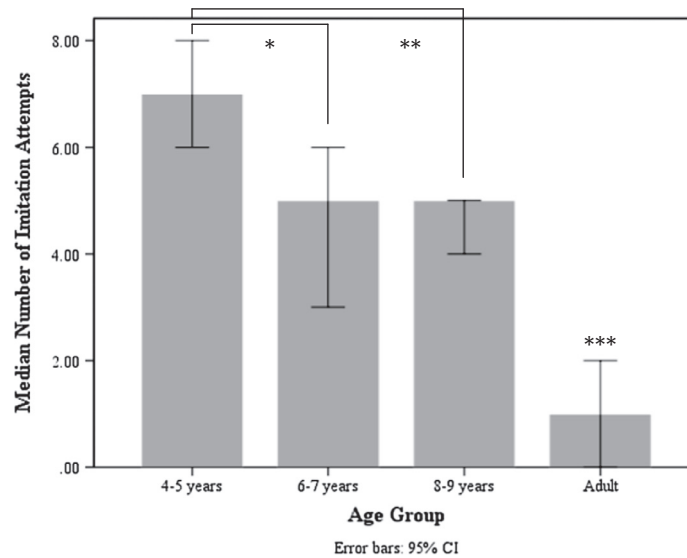


Fig. 2. Median number of 'pure' imitation attempts by age group. The asterisks above the adult bar denote that these participants were significantly different to all other age groups. * $p < .05$, ** $p < .005$, *** $p < .001$.

attempted extraction) from the exit door, was the dominant strategy used on the MMB task. This was seen in participants' first attempt trial (68% of which met the criteria for 'pure' imitation) and overall (most common strategy across attempt trials for 67% of children). As the exit door was unlocked for all participants on the first trial, excepting those in the 0% condition for whom it was always locked, the first enactment of the socially demonstrated method allowed for successful extraction.

In spite of the dominance of this imitation response, it was mediated by age ($H(2) = 8.86$, $p = .012$). The number of imitation attempts was significantly higher for the youngest age group (4–5 years, $N = 59$) when compared with 6- to 7-year-olds (Mann Whitney, $N = 78$, $U = 1758.50$, $z = -2.41$, $p = .016$) and 8- to 9-year-olds ($N = 72$, $U = 1518.00$, $z = -2.85$, $p = .004$; see Fig. 2). The latter two groups did not significantly differ ($U = 2752.50$, $z = -0.21$, $p = .83$), nor did the experimental conditions ($H(3) = 2.54$, $p = .47$). However, an effect of condition was found for adults, who produced significantly more imitation attempts in the highest (75%) efficacy condition ($Mdn = 4$, $SD = 2.88$) compared with the 0% efficacy condition ($Mdn = 0$, $SD = 0.65$): $U = 41.50$, $z = -5.01$, $p < .001$.

3.3. How did the children's behaviour deviate from the demonstrated behaviour?

To establish which component of the method (tool, access, exit) was most likely to be modified, separate scores were created for the number of novel tools (maximum 5; hook, sweep, combined fork, combined sweep, tool end), novel access points (maximum 4; end, dial, dial opposite, entry chute) and novel exits (maximum 3; end, dial, dial opposite) used across the attempt trials ('novel' denoting 'not seen' in demonstrations, and excluding repetitions).

A significant difference was found in method component modification or the 'type' of innovation (Friedman's ANOVA $\chi^2(2) = 114.94$, $p < .001$). Specifically, participants used significantly more novel tools throughout their attempt trials ($Mdn = 1$, $SD = 0.97$) than access points ($Mdn = 0$, $SD = 0.93$; Wilcoxon signed-ranks $z = -6.35$, $p < .001$) or exits ($Mdn = 0$, $SD = 0.61$; $z = -9.21$, $p < .001$). Participants also used significantly more novel access points than novel exits ($z = -5.96$, $p < .001$). These findings are further reflected in the total number of children (out of 209) who produced at least one of the different innovation types: tool

innovation ($N = 132$), access innovation ($N = 86$) and exit innovation ($N = 26$).

In contrast, adult participants ($\chi^2(2) = 13.34$, $p = .001$) used significantly more novel exits ($Mdn = 2$, $SD = 1.14$) than novel tools ($Mdn = 2$, $SD = 0.85$; $z = -2.75$, $p = .006$). There was no significant difference in the use of novel tools and novel access points ($z = -0.18$, $p = .86$), and the difference between novel exits and novel access points neared significance ($Mdn = 1$, $SD = 1.18$; $z = -2.22$, $p = .027$, $p^* = .016$). Examining age differences in children's novel exit use, the oldest children (8–9 years: $Mdn = 1$, $SD = 0.86$) were the most proficient (6–7 years: $Mdn = 1$, $SD = 0.47$; 4–5 years: $Mdn = 1$, $SD = 0.28$), although the effect is only nearing significance ($H(2) = 5.79$, $p = .055$).

3.4. Improving behaviour efficacy: The importance of exit innovation

The experimental task was designed such that exit innovations were the only way in which behaviour could be made more efficacious. Whilst modifications of the tool and access point are innovative departures from demonstrated behaviour, without modification of the exit they are of no more 'use' than the modelled method. Innovations should solve the problem at hand (Carr, Kendal, & Flynn, under revision). Unlike the exit door, the top access points of the box are always open and thus can guarantee extraction success when used as exits. It is for this reason that only rates of exit innovation were included in the following analyses, and not rates of tool or access innovation.

Of the 209 child participants within the four social experimental conditions, only 26 individuals (12.4%) produced at least one exit innovation (age group differences are reported at the end of this section). Thus, whilst 10% of children never imitated, 87.6% of children never innovated. This is in contrast to the 33 of 45 adult participants (73.3%) who did produce at least one exit innovation. The disparity between 'pure' imitation and exit innovation as adopted task strategies, across ages, can be seen in Fig. 3. Correlational analyses, using actual ages and mean number of attempts, indicated a significant negative correlation between imitation and age ($r_s(254) = -0.35$, $p < .001$) and a significant positive correlation between exit innovation and age ($r_s(254) = 0.47$, $p < .001$).

Children's exit innovations typically appeared around the fourth attempt trial out of eight (see Table 3), suggesting that

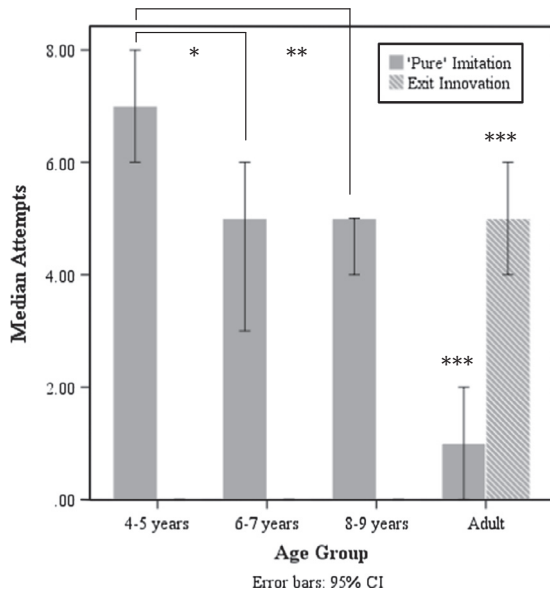


Fig. 3. Median number of 'pure' imitation and exit innovation attempts by age group. The asterisks above the adult bar denote that these participants were significantly different to all other age groups. * $p < .05$, ** $p < .005$, *** $p < .001$.

innovative problem solving was a cumulative process, with each trial or interaction with the MMB revealing more about its affordances, or that participants opted to explore once they had gained personal experience of the demonstrated method's efficacy. A clear trend of increasing exit innovation with decreasing efficacy of the demonstrated method was seen. While 23% of children in the 0% condition (where the exit door never yielded to allow extraction)

produced at least one exit innovation, this was true of only 13% of children in the 25% condition and 6% of children in the 75% and 100% conditions.

Each innovation of the 26 individuals was 'graded' according to its complexity, thereby taking into consideration the tool and access point that accompanied the new exit. Scores were as follows: (1) new exit only, (2) new exit and new tool or access point, (3) new exit, new tool and new access point. In addition to grades, innovations were also categorised by their level. Higher-level innovations were determined by their repetition (and presumed learned status), deemed to be of cultural significance given the increased likelihood of their successful transmission and acquisition by others (as opposed to an innovation that is accidental or remains in the repertoire of only one individual; see Carr, Kendal, & Flynn, under revision). A low-level innovation is defined as an 'unlearned chance innovation not repeated by the individual', to be contrasted with a mid-level 'individually learned innovation repeated by the individual' (the high-level category, 'individually learned innovation that is acquired by others', does not apply as this study did not allow for transmission of innovations to other individuals). The occurrence and number of repetitions, used to determine the level of the innovation, can be seen in the right-hand column of Table 3.

There were no significant differences in the number of exit innovation extractions by age group ($H(2) = 5.39$, $p = .07$). However, as Table 3 indicates, there were age differences when considering exit innovations more closely. Of the five exit innovators in the 4–5 age group, no one individual discovered more than one novel exit. The number of individuals doing so increased in the 6–7 group ($M = 1.43$), and again in the oldest group ($M = 2.00$). Moreover, although overall there were very few repetitions of exit innovations ($M = 0.81$), adult participants displayed a higher mean number of exit innovation repetitions ($M = 3.21$) than children

Table 3
Exit innovations: Participant and innovation characteristics.

Participant characteristics			First exit innovation trial	No. of new exit innovations	Grade	Repetitions of exit innovations
Age (Years)	Sex	Condition				
4–5	F	0%	1	1	1	4
	M	0%	8	1	1	0
	M	0%	2	1	3	0
	M	25%	6	1	1	0
	M	25%	5	1	3,1	1
6–7	F	75%	2	1	2	0
	M	0%	2	3	2,3,2,3	1
	M	0%	3	1	3	0
	M	0%	1	2	3,2	2
	M	0%	2	1	1	0
	M	25%	6	1	2	0
	M	25%	5	1	1	2
8–9	F	0%	3	2	1,2	0
	F	0%	3	2	1	1
	F	0%	8	1	1	0
	F	0%	4	1	1	2
	F	25%	4	2	1	1
	F	25%	7	1	2	0
	F	75%	3	3	3	0
	F	100%	6	1	2	0
	F	100%	4	1	1	2
	M	0%	1	2	3	2
	M	0%	4	3	3,2,3	0
	M	0%	2	3	1,2,3,1	2
	M	75%	5	2	2,1	0
	M	100%	2	3	3,3,2	1

Note. 'Grade' reflects the complexity (3 = most complex) of the novel behaviour as a whole (tool, access and exit), and are written in the order in which they were displayed. 'Repetitions of exit innovations' is a count of the number of times a newly discovered exit (i.e., not the exit door) was used again. It does not denote how many *different* exit innovations were repeated. The 'new' in 'Number of new exit innovations' relates to the child, and excludes the exit door used in social demonstrations. It does not denote how *many* capsules were extracted, only how many of the access points were discovered as exits.

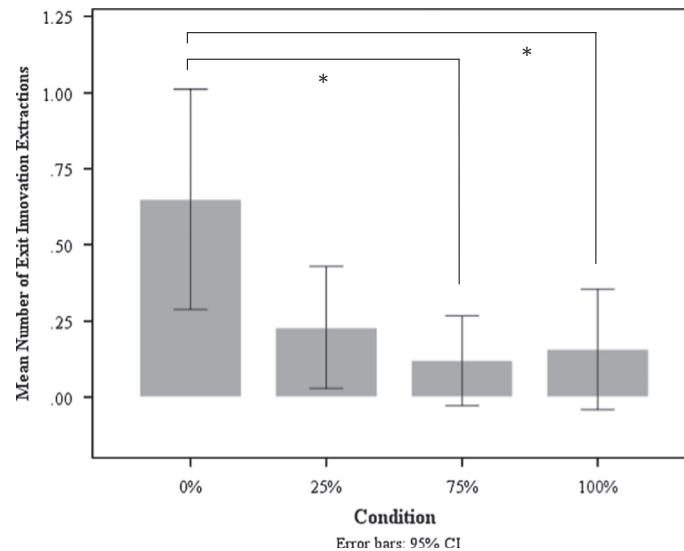


Fig. 4. Mean number of exit innovation extractions across child experimental groups. Although non-parametric statistics were conducted, the means are displayed here given that the median score for each group was 0. * $p < .05$.

($M = 0.81$), including those of the eldest children ($M = 0.79$), thus evidencing more innovations of mid-level status. A variety of 'grades' of innovation complexity were seen within each age group. While the innovations of some participants increased in complexity (progressing from a lower to higher grade during attempt trials), this trend was reversed for others.

Examining the number of exit innovations more widely across the four experimental groups (children only; Fig. 4), a significant effect of condition was found (Kruskal–Wallis $H(3) = 10.82$, $p = .01$). As it was predicted that those participants in the lowest efficacy conditions would innovate more than those in the higher efficacy conditions, a number of follow-up analyses were conducted. The results of these supported our predictions: participants in the 0% efficacy condition ($N = 60$, $Mdn = 0$, $SD = 1.40$) attained a significantly greater number of innovative extractions than participants in the 75% ($N = 50$, $Mdn = 0$, $SD = 0.52$; Mann Whitney $U = 1234.00$, $z = -2.54$, $p = .01$) and 100% conditions ($N = 51$, $Mdn = 0$, $SD = 0.70$; $U = 1261.50$, $z = -2.54$, $p = .01$), but not 25% ($N = 48$, $Mdn = 0$, $SD = 0.69$; $U = 1269.00$, $z = -1.56$, $p = .12$). The 25% condition did not significantly differ from the two higher efficacy conditions.

Of the 33 adults who produced one or more exit innovations, 23 belonged to the 0% efficacy condition and 10 to the 75% efficacy condition. Complementing the effect of condition found for children, adult participants in the 0% condition ($Mdn = 6.5$, $SD = 1.83$) attained a significantly greater number of innovative extractions than those in the 75% condition ($Mdn = 0$, $SD = 2.62$): $U = 63.00$, $z = -4.38$, $p < .001$.

In addition to group differences in the *performance* of exit innovations (including their repetition), we find differences in the *production* of exit innovations (new exit innovations only). Considering only new (to the child) exit innovations, the effect of condition was again significant (Kruskal–Wallis $H(3) = 10.63$, $p = .01$). Participants in the 0% condition produced significantly more new exit innovations across their eight attempt trials ($Mdn = 0$, $SD = 0.85$) than 75% ($Mdn = 0$, $SD = 0.52$; $U = 1243.00$, $z = -2.45$, $p = .014$) and 100% participants ($Mdn = 0$, $SD = 0.46$; $U = 1260.5$, $z = -2.54$, $p = .011$). The effect of age was nearing significance ($H(2) = 5.79$, $p = .055$) with the older age groups producing more exit innovations than the youngest group (as also suggested by Table 3).

4. Discussion

Here we addressed the question of how children of different ages trade-off social versus asocial learning based on the efficacy of an observed solution. We also considered how innovation, through modification in tool use, develops. Lower levels of observed solution efficacy were associated with increased (exit) innovation in children, with older children being more likely to innovate than younger children. Between 6–7 years and adulthood, imitation of the socially demonstrated method decreased and innovation increased. Contrary to expectation, reduced imitation in response to lower levels of solution efficacy was not found for children. It was, however, seen in adults.

4.1. Fidelity to, and deviations from, the socially demonstrated method

Children reproduced modelled behaviour with high levels of fidelity across the different efficacy conditions, supporting previous research indicating imitation is one of the major learning mechanisms used by children (Hopper et al., 2010; Horner & Whiten, 2005; Whiten & Flynn, 2010). The pervasiveness of imitation occurred in spite of permission to deviate ('try anything you like') and repetition of the goal ('see if you can get the egg out of the box'), alongside explicit linguistic cues as to whether or not the goal had been achieved. Faithful reproduction of modelled behaviour cannot be ascribed to task difficulty (known to increase imitation in children: Williamson & Meltzoff, 2011; Williamson et al., 2008), as the majority of no-demonstration control participants were able to solve the task asocially. Three possible interpretations remain.

First, children are poor at evaluating efficacy of observed information (and indeed personal information when they reproduce the socially demonstrated method). Although the exit innovation findings in the current study stem from a small number of children, meaning this interpretation cannot be completely ruled out, the significant effect of experimental condition between the 0% and 75/100% groups does not appear to support the notion that children are poor at evaluating efficacy, nor do findings of prior research (Pasquini et al., 2007). Second, contradicting the actions of an adult demonstrator, by opting not to reproduce demonstrated

behaviour, is an unfavourable option for children (due to adults' general level of perceived competence, Wood et al., 2012, or their modelling of normative behaviour). Yet previous evidence suggests that when there is sufficient reason to do so (i.e., the model is unreliable, actions are accidental, and behaviour is inefficacious), children will deviate (Birch et al., 2008; Carpenter et al., 1998; Williamson et al., 2008; Zmyj, Buttelmann, Carpenter, & Daum, 2010). Moreover, children were seen to deviate from the adult demonstrator, principally by trying out different tools (but less so the crucial exits). The third and final interpretation is that generating novel behaviour (as an alternative to imitation), capable of successfully altering the outcome of the task, was cognitively demanding following social demonstrations and that either the capacity or motivation to do so was lacking. Whilst not mutually exclusive, we propose that the competence interpretation (solely or in combination with a normative explanation given below) best explains the current findings – especially as the ability to use (innovate) a new exit increased from 8–9 years into adulthood – and aligns with previous research (children: Beck et al., 2011; callitrichids: Kendal, Coe, & Laland, 2005).

A number of important developmental progressions were uncovered in the present study, and suggest that reliance on social learning mechanisms is in part determined by age. In spite of the dominant imitation response, age effects were found regarding fidelity: 4- to 5-year-olds demonstrated more faithful imitation (enacting this strategy across more attempt trials) than older children. Imitation fidelity continued to decrease into adulthood. In the context of children's novel puzzle box interactions, wherein there is an explicit goal (tasks are not causally opaque), imitation thus appears to increase between the ages of 3 and 5 years (Flynn & Whiten, 2008; McGuigan, Whiten, Flynn, & Horner, 2007) before plateauing around the age of 6 years (present study). Consistent with Rakoczy, Hamann, Warneken, and Tomasello's (2010) observation of children deeming adults' demonstrated behaviour to be normatively correct, several children remarked, following demonstrations, 'so that's how you play the game'. This indication of rule learning or convention acquisition, together with children's general reluctance to depart from demonstrated behaviour, suggests normativity had a part to play in the findings. The age-driven decline in imitation could be facilitated by an age-driven decline in normativity and, relatedly, conformity (Walker & Andrade, 1996). Conformity also appears to be reduced for children, from the age of four years, when making judgements in more objective and less socially arbitrary domains (judging object *functions* as opposed to object *labels*; Schillaci & Kelemen, 2014).

4.2. Exit innovation: The rate and influence of observed behaviour efficacy

In line with cultural evolution theory (Boyd & Richerson, 1985; Richerson & Boyd, 2005) and previous experimental studies (Whiten & Flynn, 2010), in the current study a small minority of innovators emerged from a large population of 'followers'. Exit innovations were produced by only 26 of 209 children following social demonstrations. The majority of children failed to recognise that exit innovations represented the sole way in which behaviour could be made more efficacious, such that a focus on behavioural *means* (tools used) as opposed to behavioural *outcome* prevailed (it could also be that the tools were highly salient to the children, by being the first object that was selected by the demonstrator, but less so to the more experienced adults). Those who continued with the socially demonstrated method when it was never efficacious (0% condition) or rarely efficacious (25% condition) may have found the social affiliation function of imitation (Over & Carpenter, 2012) rewarding. In future, it would be interesting to introduce a

competition element whereby children would be encouraged to gain more stickers than the demonstrator.

Functional fixedness is a unique challenge for artefact tasks, and may account for the rarity of innovation. It describes a phenomenon whereby an object's known conventional function prevents an appreciation of its alternative uses (German & Defeyter, 2000); in the case of the MMB, the top access points conventionally function as tool entrances, not capsule exits. The somewhat counter-intuitive developmental trend of functional fixedness (affecting 7-year-olds to a greater extent than 5-year-olds; Defeyter, Avons, & German, 2007) likely impedes the emergence of innovation; hampering its production just at the time that increasing cognitive flexibility may better enable it. Although innovation increased with age, exit innovators were nevertheless very rare amongst even the oldest child age group. Executive functions may have a similar limiting effect. Inhibitory control skills develop significantly in the preschool years, but children do not show mature or advanced levels of performance in some executive abilities until aged 9–10 years or above (e.g., action inhibition: Simpson, Cooper, Gillmeister, & Riggs, 2013; planning: Tecwyn, Thorpe, & Chappell, 2014). Together with the late-developing inductive reasoning, permitting a variety of inferences to be made about a single item that fits multiple categories (Bright & Feeney, 2014), inhibiting prepotent responses and considering multiple possible outcomes prior to action will surely better enable innovation. With age, our participants became less restricted in their exit innovation capabilities – perhaps indicating the requirement for mature executive functions and more general cognitive maturity and flexibility to overcome the functional fixedness obstacle.

Rates of exit innovation increased from 8–9 years and were influenced by observed behaviour efficacy. Participants who experienced the lowest level of solution efficacy (the exit door was always locked) produced a greater number of innovative extractions than participants with a 75% or 100% level of observed solution efficacy. These latter two conditions were arguably the least conducive to innovation since they provided participants with a solution that always, or nearly always, worked. Yet, innovation is not just about solving a problem but exploring the world also. Indeed, Wood et al. (2013a) discovered that children are motivated to acquire multiple solutions to a problem even without the potential of a greater reward. In the current study, the 75% and 100% participants could plausibly afford to explore *more* than the 25% or 0% participants in the knowledge that they already have a functional method in their repertoire, meaning potentially better ways of accomplishing the goal could be sought. It may be that children's performance was influenced by an adult model-based bias (a puppet was used for demonstrations in Wood et al.), but an alternative interpretation is suggested by the adult findings. Of the 12 adult participants who did not produce an exit innovation, 11 belonged to the 75% condition. Given that adults are not cognitively constrained in the same manner as children, it appears that they deduced no necessity in deviating from the socially demonstrated method when it was largely functional. Therefore, our results show it was likely necessity, not opportunity (implicated in the innovative tool use of various non-human primate species; Koops, Visalberghi, & van Schaik, 2014), that drove participants to innovate.

Individual learning from performing (exit) innovations was evidenced in two ways: repetition of an innovation, and/or production of more than one innovation. According to the former criteria, 10 of the 26 child exit innovators produced low-level innovations where there was no evidence of learning (note, however, that two of these individuals produced an exit innovation on the eighth trial, preventing subsequent assessment of learning). The greater propensity for innovation repetitions in adults (only 2 out of 33

producing low-level innovations) hints at the operation of more sophisticated learning and executive processes, but also at a potential disparity in approach to the MMB task. It is possible that adults were more goal-directed, prioritising the extraction of the capsule over the attempt trials, whereas children, when they chose to deviate from the social method and explore, did so in a more playful and 'random' manner. This is supported also by the varying complexities, or 'grades', of innovations when they were produced by the children. As an aside, in the context of the MMB task we cannot necessarily ascribe greater theoretical significance to exit innovations that were accompanied by a novel tool and/or access point: the latter do not improve behaviour efficacy. However, in other contexts, innovation across all components may be regarded as more insightful. Returning to the exit innovation findings, given children's capacity to incorporate newly presented task solutions into their behavioural repertoires (Wood et al., 2013a), we propose that it was the generation of alternative solutions as opposed to the switching between them which created difficulties with the current task.

4.3. Implications

Comparing our findings with those of Beck et al. (2011), where children succeeded at a novel hook invention task around 8 years of age, we provisionally suggest that innovation by modification and innovation by novel invention have somewhat distinct developmental trajectories. However, this can only be confirmed with further research including a variety of tasks. We also posit that, whilst innovation of any form is made challenging by a lack of certain cognitive abilities (particularly higher-level executive functions), individuals attempting innovation by modification are especially vulnerable to a canalisation or conservatism effect of prior social demonstrations. This is manifest in functional fixedness in tool-use tasks. Whilst the indication is that the task was more difficult in the absence of prior social demonstrations (fewer capsules were extracted in the no-demonstration than 100% condition), without these prior demonstrations participants were more exploratory and attempted a greater number of alternative task methods. Wood et al. (2013a) and Bonawitz et al. (2011) have similarly found observation and pedagogy to lead to restricted exploration and learning. The cost of quick and 'cheap' social information acquisition is ultimately behavioural canalisation: becoming stuck on a particular method, and in turn blind to potential alternatives. Reducing the social context in experiments, to ascertain the extent to which innovation is inhibited by pedagogy, remains an imperative objective.

Laland (2004, p. 11) speculated that, 'If innovation is risky and associated with costs, then it is likely to be employed as a last resort... when socially learned strategies have proven unproductive'. Though there was no indication that innovation would be risky or costly for those in the low efficacy conditions of the current study (when the socially demonstrated method was unsuccessful and the reward could not be retrieved), our findings of rare and limited innovation, even in older and more competent individuals, do indeed suggest that children employ innovation as a last resort.

Acknowledgements

This work was supported by an ESRC studentship awarded to K.C. and an ESRC Grant awarded to E.G.F. We would like to thank the children, schools and students who participated in this study, Andrew Long for assistance with the design and manufacture of the task, and Natalia Dutra for inter-rater coding.

References

- Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition*, 119, 301–306. <http://dx.doi.org/10.1016/j.cognition.2011.01.003>.
- Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology*, 53A, 153–164.
- Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, 107, 1018–1034. <http://dx.doi.org/10.1016/j.cognition.2007.12.008>.
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120, 322–330. <http://dx.doi.org/10.1016/j.cognition.2010.10.001>.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, US: University of Chicago Press.
- Bright, A. K., & Feeney, A. (2014). Causal knowledge and the development of inductive reasoning. *Journal of Experimental Child Psychology*, 122, 48–61. <http://dx.doi.org/10.1016/j.jecp.2013.11.015>.
- Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior & Development*, 21, 315–330.
- Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, 8, 13–20.
- Cutting, N., Apperly, I. A., Chappell, J., & Beck, S. R. (2014). The puzzling difficulty of tool innovation: Why can't children piece their knowledge together? *Journal of Experimental Child Psychology*, 125, 110–117. <http://dx.doi.org/10.1016/j.jecp.2013.11.010>.
- Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science*, 335, 1114–1118. <http://dx.doi.org/10.1126/science.1213969>.
- Defeyter, M. A., Avons, S. E., & German, T. C. (2007). Developmental changes in information central to artifact representation: Evidence from 'functional fluency' tasks. *Developmental Science*, 10, 538–546. <http://dx.doi.org/10.1111/j.1467-7687.2007.00617.x>.
- Flynn, E. (2008). Investigating children as cultural magnets: Do young children transmit redundant information along diffusion chains? *Philosophical Transactions of the Royal Society B*, 363, 3541–3551. <http://dx.doi.org/10.1098/rstb.2008.0136>.
- Flynn, E., & Smith, K. (2012). Investigating the mechanisms of cultural acquisition: How pervasive is overimitation in adults? *Social Psychology*, 43, 185–195. <http://dx.doi.org/10.1027/1864-9335/a000119>.
- Flynn, E., & Whiten, A. (2008). Cultural transmission of tool use in young children: A diffusion chain study. *Social Development*, 17, 699–718. <http://dx.doi.org/10.1111/j.1467-9507.2007.00453.x>.
- Galef, B. G., Jr., & Laland, K. N. (2005). Social learning in animals: Empirical studies and theoretical models. *BioScience*, 55, 489–499.
- Gardiner, A. K., Bjorklund, D. F., Greif, M. L., & Gray, S. K. (2012). Choosing and using tools: Prior experience and task difficulty influence preschoolers' tool-use strategies. *Cognitive Development*, 27, 240–254. <http://dx.doi.org/10.1016/j.cogdev.2012.05.001>.
- German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review*, 7, 707–712. <http://dx.doi.org/10.3758/BF03213010>.
- Hanus, D., Menden, N., Tennie, C., & Call, J. (2011). Comparing the performances of apes (*Gorilla gorilla*, *Pan troglodytes*, *Pongo pygmaeus*) and human children (*Homo sapiens*) in the floating peanut task. *PLoS ONE*, 6, e19555. <http://dx.doi.org/10.1371/journal.pone.0019555>.
- Hopper, L. M., Flynn, E. G., Wood, L. A. N., & Whiten, A. (2010). Observational learning of tool use in children: Investigating cultural spread through diffusion chains and learning mechanisms through ghost displays. *Journal of Experimental Child Psychology*, 106, 82–97. <http://dx.doi.org/10.1016/j.jecp.2009.12.001>.
- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, 8, 164–181. <http://dx.doi.org/10.1007/s10071-004-0239-6>.
- Kameda, T., & Nakanishi, D. (2002). Cost-benefit analysis of social/cultural learning in a nonstationary uncertain environment: An evolutionary simulation and an experiment with human subjects. *Evolution and Human Behavior*, 23, 373–393.
- Kendal, R. L., Coe, R. L., & Laland, K. N. (2005). Age differences in neophilia, exploration, and innovation in family groups of callitrichid monkeys. *American Journal of Primatology*, 66, 167–188. <http://dx.doi.org/10.1002/ajp.20136>.
- Kendal, R. L., Coolen, I., & Laland, K. N. (2009). Adaptive trade-offs in the use of social and personal information. In R. Dukas & J. M. Ratcliffe (Eds.), *Cognitive ecology II* (pp. 249–271). Chicago, US: University of Chicago Press.
- Kendal, R. L., Coolen, I., van Bergen, Y., & Laland, K. N. (2005). Trade-offs in the adaptive use of social and asocial learning. *Advances in the Study of Behaviour*, 35, 333–379.
- Koenig, M. A., & Sabbagh, M. A. (2013). Selective social learning: New perspectives on learning from others. *Developmental Psychology*, 49, 399–403. <http://dx.doi.org/10.1037/a0031619>.
- Koops, K., Visalberghi, E., & van Schaik, C. P. (2014). The ecology of primate material culture. *Biology Letters*, 10, 20140508. <http://dx.doi.org/10.1098/rsbl.2014.0508>.
- Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, 32, 4–14.

- Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B*, 367, 2171–2180. <http://dx.doi.org/10.1098/rstb.2012.0119>.
- Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, 104, 19751–19756. <http://dx.doi.org/10.1073/pnas.0704452104>.
- Matheson, H., Moore, C., & Akhtar, N. (2013). The development of social learning in interactive and observational contexts. *Journal of Experimental Child Psychology*, 114, 161–172. <http://dx.doi.org/10.1016/j.jecp.2012.09.003>.
- McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development*, 22, 353–364. <http://dx.doi.org/10.1016/j.cogdev.2007.01.001>.
- Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, 49, 404–418. <http://dx.doi.org/10.1037/a0029500>.
- Nielsen, M. (2013). Young children's imitative and innovative behavior on the floating object task. *Infant and Child Development*, 22, 44–52. <http://dx.doi.org/10.1002/icd.1765>.
- Over, H., & Carpenter, M. (2012). Putting the social into social learning: Explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, 126, 182–192. <http://dx.doi.org/10.1037/a0024555>.
- Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, 43, 1216–1226. <http://dx.doi.org/10.1037/0012-1649.43.5.1216>.
- Pinkham, A. M., & Jaswal, V. K. (2011). Watch and learn? Infants privilege efficiency over pedagogy during imitative learning. *Infancy*, 16, 535–544. <http://dx.doi.org/10.1111/j.1532-7078.2010.00059.x>.
- Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better-young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology*, 28, 785–798. <http://dx.doi.org/10.1348/026151009X479178>.
- Reader, S. M., & Laland, K. N. (2003). Animal innovation: An introduction. In S. M. Reader & K. N. Laland (Eds.), *Animal Innovation* (pp. 3–35). Oxford, UK: Oxford University Press.
- Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, 15, 68–76. <http://dx.doi.org/10.1016/j.tics.2010.12.002>.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, US: University of Chicago Press.
- Rieucau, G., & Giraldeau, L.-A. (2011). Exploring the costs and benefits of social information use: An appraisal of current experimental evidence. *Philosophical Transactions of the Royal Society B*, 366, 949–957. <http://dx.doi.org/10.1098/rstb.2010.0325>.
- Schillaci, R. S., & Kelemen, D. (2014). Children's conformity when acquiring novel conventions: The case of artifacts. *Journal of Cognition and Development*, 15, 569–583. <http://dx.doi.org/10.1080/15248372.2013.784973>.
- Schulz, L. E., Hoopell, C., & Jenkins, A. C. (2008). Judicious imitation: Children differentially imitate deterministically and probabilistically effective actions. *Child Development*, 79, 395–410. <http://dx.doi.org/10.1111/j.1467-8624.2007.01132.x>.
- Simpson, A., Cooper, N. R., Gillmeister, H., & Riggs, K. J. (2013). Seeing triggers acting, hearing does not trigger saying: Evidence from children's weak inhibition. *Cognition*, 128, 103–112. <http://dx.doi.org/10.1016/j.cognition.2013.03.015>.
- Tecwyn, E. C., Thorpe, S. K. S., & Chappell, J. (2014). Development of planning in 4- to 10-year-old children: Reducing inhibitory demands does not improve performance. *Journal of Experimental Child Psychology*, 125, 85–101. <http://dx.doi.org/10.1016/j.jecp.2014.02.006>.
- Tomasello, M. (1990). Cultural transmission in the tool use and communicatory signaling of chimpanzees? In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes* (pp. 274–311). Cambridge, UK: Cambridge University Press.
- van Bergen, Y., Coolen, I., & Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society B*, 271, 957–962. <http://dx.doi.org/10.1098/rspb.2004.2684>.
- Walker, M. B., & Andrade, M. G. (1996). Conformity in the Asch task as a function of age. *The Journal of Social Psychology*, 136, 367–372.
- Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development*, 72, 431–443. <http://dx.doi.org/10.1111/1467-8624.00288>.
- Whiten, A., Custance, D. M., Gomez, J., Teixidor, P., & Bard, K. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 110, 3–14. <http://dx.doi.org/10.1037/0735-7036.110.1.3>.
- Whiten, A., & Flynn, E. (2010). The transmission and evolution of experimental microcultures in groups of young children. *Developmental Psychology*, 46, 1694–1709. <http://dx.doi.org/10.1037/a0020786>.
- Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B*, 364, 2417–2428. <http://dx.doi.org/10.1098/rstb.2009.0069>.
- Williamson, R. A., & Meltzoff, A. (2011). Own and others' prior experiences influence children's imitation of causal acts. *Cognitive Development*, 26, 260–268. <http://dx.doi.org/10.1016/j.cogdev.2011.04.002>.
- Williamson, R. A., Meltzoff, A. N., & Markman, E. M. (2008). Prior experiences and perceived efficacy influence 3-year-olds' imitation. *Developmental Psychology*, 44, 275–285. <http://dx.doi.org/10.1037/0012-1649.44.1.275>.
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2012). Context-dependent model-based biases in cultural transmission: Children's imitation is affected by model age over model knowledge state. *Evolution and Human Behavior*, 33, 387–394. <http://dx.doi.org/10.1016/j.evolhumbehav.2011.11.010>.
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013a). Copy me or copy you? The effect of prior experience on social learning. *Cognition*, 127, 203–213. <http://dx.doi.org/10.1016/j.cognition.2013.01.002>.
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013b). Whom do children copy? Model-based biases in social learning. *Developmental Review*, 33, 341–356. <http://dx.doi.org/10.1016/j.dr.2013.08.002>.
- Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology*, 106, 208–220. <http://dx.doi.org/10.1016/j.jecp.2010.03.002>.